

Strietholt, Rolf; Terhart, Ewald

Referendare beurteilen. Eine explorative Analyse von Beurteilungsinstrumenten in der Zweiten Phase der Lehrerbildung

Zeitschrift für Pädagogik 55 (2009) 4, S. 622-645



Quellenangabe/ Reference:

Strietholt, Rolf; Terhart, Ewald: Referendare beurteilen. Eine explorative Analyse von Beurteilungsinstrumenten in der Zweiten Phase der Lehrerbildung - In: Zeitschrift für Pädagogik 55 (2009) 4, S. 622-645 - URN: urn:nbn:de:0111-opus-42666 - DOI: 10.25656/01:4266

<https://nbn-resolving.org/urn:nbn:de:0111-opus-42666>

<https://doi.org/10.25656/01:4266>

in Kooperation mit / in cooperation with:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipt.de
Internet: www.pedocs.de

Inhaltsverzeichnis

Thementeil: Werte – Die Rückkehr zur Erziehung

Hartmut von Hentig

Das Ethos der Erziehung. Was ist in ihr elementar? 509

Gertrud Nunner-Winkler

Prozesse moralischen Lernens und Entlernens 528

Matthias Gronemeyer

Wie kommt das Böse in den Menschen – und wie wieder heraus? Ein
philosophischer Kommentar zu Serkan A. und Spyridon L. 549

Margit Stein

Werteerziehungsansätze an weiterführenden Schulen in ihrem Zusammenhang
mit strukturellen Schulbedingungen 562

Thomas Gensicke

Jugendlicher Zeitgeist und Wertewandel 580

Allgemeiner Teil

Esther Dominique Klein/Svenja Mareike Kühn/Isabell van Ackeren/Rainer Block

Wie zentral sind zentrale Prüfungen? – Abschlussprüfungen am Ende der
Sekundarstufe II im nationalen und internationalen Vergleich 596

Rolf Strietholt/Ewald Terhart

Referendare beurteilen. Eine explorative Analyse von Beurteilungsinstrumenten
in der Zweiten Phase der Lehrerbildung 622

Besprechungen

Marcelo Caruso

Colin Brock/Lila Zia Levers (Hrsg.): Aspects of Education in the Middle East and North Africa 646

Heinz-Elmar Tenorth

Wolfgang Brezinka: Pädagogik in Österreich. Die Geschichte des Faches an den Universitäten vom 18. bis zum 21. Jahrhundert 648

Wolfgang Jantzen

Dagmar Hänsel: Die NS-Zeit als Gewinn für Hilfsschullehrer.

Dagmar Hänsel: Karl Tornow als Wegbereiter der sonderpädagogischen Profession. Die Grundlegung des Bestehenden in der NS-Zeit 650

Hanno Schmitt

Sieglinde Luise Ellger-Rüttgardt: Geschichte der Sonderpädagogik. Eine Einführung 653

Melanie Fabel-Lamla

Fritz Reheis: Bildung contra Turboschule! Ein Plädoyer.

Wiebke Lohfeld (Hrsg.): Gute Schulen in schlechter Gesellschaft.

Hans-Peter Gerstner/Martin Wetz: Einführung in die Theorie der Schule.

Gregor Lang-Wojtasik: Schule in der Weltgesellschaft. Herausforderungen und Perspektiven einer Schultheorie jenseits der Moderne 655

Cristina Allemann-Ghionda

Pedro Rosselló: La teoría de las corrientes educativas y otros ensayos

Christel Adick: Vergleichende Erziehungswissenschaft

Doris Edelmann: Pädagogische Professionalität im transnationalen Raum 660

Dokumentation

Pädagogische Neuerscheinungen 665

Referendare beurteilen. Eine explorative Analyse von Beurteilungsinstrumenten in der Zweiten Phase der Lehrerbildung

Zusammenfassung: Die Frage nach den Möglichkeiten einer zuverlässigen Beurteilung der beruflichen Kompetenzen von Lehrerinnen und Lehrern wird innerhalb der wissenschaftlichen Fachdiskussion, aber auch in den Lehrerverbänden und in der Öffentlichkeit kontrovers diskutiert. Dabei wird häufig übersehen, dass die Beurteilung von (angehenden) Lehrkräften seit Jahrzehnten stattfindet: im Rahmen der Zweiten Ausbildungsphase, dem Referendariat oder Vorbereitungsdienst. Nach einer kurzen Übersicht über neuere Forschungsergebnisse zur Situation der Zweiten Phase wird über ein Forschungsprojekt berichtet, in dem die in der Zweiten Phase von Ausbildern (Fachleitern, Hauptseminarleitern) verwendeten Beurteilungsschemata und -manuale analysiert worden sind. Auf der Basis von 201 Beurteilungsschemata und 87 zusätzlicher Fragebögen zur Entstehung und zur Verwendung der Schemata wird eine Übersicht über die Praxis des Beurteilens im Referendariat gegeben. Die Ergebnisse machen u.a. deutlich, dass dem Großteil der Instrumente ein auf Unterrichten konzentriertes Bild der Lehrertätigkeit zugrunde liegt, die Breite der Standards für die Lehrerbildung also nicht abgebildet wird. In jedem zweiten Instrument bleibt unklar, wie der Beurteilungsakt im Einzelnen vollzogen wird. Die testdiagnostische Qualität der Instrumente ist in keinem Fall geprüft worden. Insgesamt wird deutlich, dass hinsichtlich der Qualität der Beurteilungsprozesse und -instrumente, die in der Zweiten Phase der Lehrerbildung eingesetzt werden, ein dringender Entwicklungsbedarf besteht.

1. Einleitung

Der Ausbaustand der empirischen Forschung zur Lehrerbildung entspricht noch keineswegs der Bedeutung, die der Lehrerbildung für die Qualität des Schulsystems zukommt. Obwohl der Umfang der empirischen Forschung zur Lehrerbildung zugenommen hat und in einer Reihe von Publikationen dokumentiert worden ist (vgl. Lemmermöhle/Jareis 2003; Blömeke u.a. 2004; Allemann-Ghionda/Terhart 2006; Hilligus/Rinkens 2006; Seifried/Abel 2006; Lüders/Wissinger 2007; Lemmermöhle u.a. 2007; Blömeke u.a. 2008; Frey 2008; Rothland/Terhart 2009), können sehr viele Entscheidungen über die Gestaltung von Strukturen, Prozessen und Inhalten der Lehrerbildung nicht oder nur zu einem geringen Teil durch Rekurs auf aussagekräftige Forschungsergebnisse getroffen werden. Eine tatsächlich „evidenzbasierte“ Gestaltung der Lehrerbildung steht noch in weiter Ferne. Dieses Forschungsdefizit betrifft makro- wie mikroskopische Probleme der Lehrerbildung, und es betrifft – legt man eine andere Systematisierung zugrunde – deren gesamten Verlauf: die erste, universitäre Phase, die zweite Phase, also den Vorbereitungsdienst bzw. das Referendariat und schließlich die mittlerweile so genannte dritte Phase: das kontinuierliche Weiterlernen von Lehrerinnen und Lehrern im Beruf.

Im Folgenden wird über die Ergebnisse eines explorativ angelegten Forschungsprojekts berichtet, in dessen Rahmen solche Beurteilungsschemata etc. analysiert wurden, die von Ausbildern in der Zweiten Phase im Rahmen der Beurteilung von Referendaren eingesetzt werden.¹ Damit wird nicht nur dieses spezielle Element der Lehrerbildung durch Forschung erschlossen, sondern auch ein Beitrag zur Diskussion um „Standards für die Lehrerbildung“ geleistet. Es geht im Folgenden allerdings nicht um die „normative“ Ausfaltung und Ausdifferenzierung von wünschenswerten Kompetenzen und idealen Standards, sondern um diejenigen Instrumente, die gegenwärtig *in der laufenden Praxis* der zweiten Phase der Lehrerbildung zur Beurteilung von Referendaren bereits eingesetzt werden, um den Charakter dieser Instrumente sowie um die in diese Instrumente eingelagerten Leitbilder und Bewertungskriterien.

Bevor jedoch über dieses Projekt und einige seiner zentralen Ergebnisse berichtet wird, soll zunächst eine knappe Übersicht über aktuelle Forschungsergebnisse zur Zweiten Phase gegeben werden.

2. Neuere Forschungsergebnisse zum Referendariat

Im Referendariat findet nach langen Jahren des Studiums eine erste, ebenso konzentrierte wie kontinuierliche Auseinandersetzung mit dem beruflichen Praxisfeld Schule und Unterricht statt.² Die Beurteilung des Referendariats durch die Referendare selbst ist deutlich ambivalent: Einerseits wird der allmähliche Aufbau von berufspraktischer Kompetenz – gerade auch in Absetzung zur universitären Phase – *positiv* gewürdigt, wobei die fachbezogenen Seminare deutlich positiver bewertet werden als die allgemein-didaktisch bzw. schulpädagogisch ausgerichteten Hauptseminare. Andererseits erscheinen nach der vergleichsweise freien universitären Lebensphase, mit ihrer Aneignung von Fachwissen und der Beschäftigung mit pädagogischen Inhalten, die unmittelbare Auseinandersetzung mit und das Einleben in die „echte Realität“ der Schule als *sehr belastend* (vgl. die Studie von Schubarth u.a. 2007 in Brandenburg; zur besonderen emotionalen Belastetheit von Referendaren im Vergleich mit Lehramtsstudierenden und Lehrkräften vgl. Weimar 2005, S. 211; plakativ Katzenbach 1999). Vom Studium geprägte Vorstellungen und Erwartungen darüber, wie viel man vom Fachwissen und von pädagogisch-didaktischen Konzeptionen in die Praxis umsetzen kann, aber auch manche grundlegende pädagogisch-didaktische Ambition werden durch unmittelbare Erfah-

-
- 1 Die Untersuchung wurde u.a. durch Mittel des Eidgenössischen Hochschulinstituts für Berufsbildung (EHB) gefördert und an der Universität Münster durchgeführt. Dass es in dem Projekt um die Analyse von Beurteilungsinstrumenten aus der Zweiten Phase der Lehrerbildung geht, darf keinesfalls als Hinweis darauf gedeutet werden, dass die Verfasser die Beurteilungspraxis in der Ersten (universitären) Phase – etwa bei der ersten Staatsprüfung – für nicht explorationsbedürftig halten. Im Gegenteil.
 - 2 Zur Situation der zweiten Phase generell sowie in einzelnen Bundesländern vgl. Walke 2007; Bötting/Thomas 2007; Böhner 2008; zur Situation des Referendariats in der Ausbildung für Lehrkräfte an beruflichen Schulen bzw. Berufskollegs vgl. Strecke 2007.

rungen mit real existierenden Schülern, Kollegen und Eltern deutlich relativiert – eine Erfahrung, die unter dem Stichwort „Praxischock“ seit Jahrzehnten in der Fachliteratur zur beruflichen Sozialisation nicht nur im Lehrerberuf, sondern auch in anderen akademischen Berufen erörtert wird.

Döbrich/Abs (2006, 2007, 2008) führen am DIPF eine Vollerhebung (Befragung) bei Ausbildern, Referendaren und Mentoren in Hessen durch. Ausbilder beurteilen die Situation und die institutionelle Leistung der Zweiten Phase durchweg besser als Referendare und Mentoren. Referendare sehen für sich die deutlichsten Kompetenzsteigerungen im Bereich des Unterrichtens und die geringsten Kompetenzsteigerungen im Bereich der Elternarbeit, des Umgangs mit kulturellen Unterschieden ihrer Schüler und der Schulentwicklung. Dieser auf Selbsteinschätzungen basierende Befund zur unterrichtszentrierten Kompetenzentwicklung wird durch ein analoges Ergebnis der Studie von Frey (2008, S. 210 bzw. 220) bestätigt. Die Verteilung der Antworten zu einem speziellen Problem – Werden Referendare nach einheitlichen Kriterien beurteilt? – soll an dieser Stelle ausführlicher dokumentiert werden. Auch bei diesem Problem zeigt sich die typische Perspektivendifferenz zwischen den drei befragten Beteiligungsgruppen.

	Aussage: „Die Bewertung von Referendaren folgt einheitlichen Kriterien“				
	trifft gar nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll zu	
Ausbilder	2,5	27,2	60,2	10,1	Prozent
Referendare	19,8	41,8	33,6	4,7	Prozent
Mentoren	14,3	47,4	35,8	2,5	Prozent

Abb. 1: Einheitlichkeit der Beurteilung von Referendaren (Prozentangaben; Döbrich/Abs 2008)

Eine in Nordrhein-Westfalen durchgeführte standardisierte Befragung von Hauptseminarleitern (das sind die für den allgemein-didaktischen und schulpädagogischen Bereich zuständigen Ausbilder; 60 auswertbare Fragebögen) der Zweiten Phase ergab u.a., dass aus deren Sicht über 50 Prozent der Absolventen ungenügende bzw. begrenzte erziehungswissenschaftliche Kenntnisse in das Referendariat mitbringen (vgl. Reintjes 2007). Im Rahmen einer international-vergleichend angelegten empirischen Studie (Befragung bzw. Selbsteinschätzung von 2.150 Lehramtsstudierenden und Referendaren aus Österreich, Schweiz, Deutschland, Polen und Italien) zeigte sich, dass ein Kompetenzzuwachs im Laufe der Lehrerbildung ausschließlich bei der Sozial- und Methodenkompetenz angegeben wurde. Bei der personalen Kompetenz gab es uneinheitliche Befunde, und bei der Fachkompetenz zeigte sich, dass diese im Zuge des Wechsels von der Universität zum Studienseminar nach Auffassung der Befragten sinkt (Frey 2008, S. 223).

Folgendes Zitat aus der Studie von Englert u.a. (2006), die auf 36 Interviews und 473 Fragebögen basiert, verdeutlicht zusammenfassend die Erfahrung der Referendare während des Vorbereitungsdienstes:

„Strukturell deuten unsere Befunde auf einen doppelten Konflikt, in dem die Studienseminare stehen. Zwischen förderlicher Lernhilfe und selektierender Benotung sollen sie Referendar/innen einerseits auf dem Weg zum Lehrerberuf unterstützen, was Offenheit und Vertrauen voraussetzt, und diese andererseits begutachten, was Distanz schafft und taktisches Verhalten nahelegt. Zwischen Theorie und Praxis sollen sie einerseits grundsätzliches, fallübergreifendes Wissen und Reflektieren fruchtbar machen, andererseits aber den Alltag von Schule und Unterricht im Blick behalten. Aus Sicht der Lehramtsanwärter/innen gelingt es den Studienseminaren derzeit nur beschränkt, die angeführten Strukturkonflikte auszubalancieren. Anstatt dass Unterstützung und Benotung, Theorie und Praxis je gleichgewichtig nebeneinander zur Geltung kommen, tendieren die Referendar/innen dazu, das Studienseminar primär als Instrument der Überprüfung und als Anwalt alltagsfremder Theorie wahrzunehmen.“ (ebd., S. 466)

Trotz dieser Ambivalenzen und Widersprüche wird von den Referendaren das Referendariat insgesamt als „ertragreiche Initiation im belastenden Ausnahmezustand“ (ebd., S. 471) betrachtet: Referendar zu sein wird als schwierige biographische Sondersituation erlebt; der Nutzen des Referendariats wird jedoch nicht bezweifelt.

Symptomatisch für Belastungen während des Referendariats sind die vorgeschriebenen „Lehrproben“: Diese werden teilweise vor anderen (Mit-)Referendaren durchgeführt (Gruppenhospitation), oder aber im Beisein des Fachleiters, dem die schriftliche Unterrichtsvorbereitung vorher zur Kenntnis zu geben ist. Im Anschluss an den Unterrichtsbesuch wird eine auswertende Nachbesprechung durchgeführt. Vielfach werden diese Lehrproben von den Lehramtsanwärtern als „unecht“ empfunden; die Lehrproben selbst wie auch die anschließenden Nachbesprechungen wirken stark ritualisiert (vgl. Hoppenworth 1993). Der inszenierte Charakter ist allen Betroffenen bewusst (die Schüler eingeschlossen) und wird gleichwohl mit großem Ernst durchgehalten. Denn schließlich hat die „unechte“ Lehrprobe eine „echte“, d.h. zukunftsbedeutsame Zwischennote zur Folge. Bei solchen Unterrichtsbesuchen werden dann von den Ausbildern Beurteilungsschemata, Kriterienlisten, Schätzskaalen etc. eingesetzt, die den Beurteilungsprozess in gewisser Weise systematisieren und dem Anspruch nach objektivieren sollen. Solche Beurteilungsschemata können sehr unterschiedliche Qualität haben; hierauf wird im Folgenden ausführlich eingegangen werden.

Abschließend sei auf zwei in der Literatur wiederkehrende Aspekte aufmerksam gemacht: Erstens scheinen Referendare die Abhängigkeit vom Urteil der Ausbilder als besonders negativ zu empfinden (vgl. dazu fallanalytisch Wernet 2006), ebenso die Situation der Konkurrenz um gute Noten und um die damit verbundenen Einstellungschancen. Und zweitens wird in vielen Analysen zur Zweiten Phase auf das besondere Problem der Rekrutierung und Qualifizierung der Ausbilder für ihre Tätigkeit hingewiesen (vgl. Walke 2004).

3. Beurteilungsinstrumente in der zweiten Phase

3.1 Kontext und Materialbasis

Unabhängig von der Tatsache, dass es ein allgemein einsetzbares diagnostisches Instrumentarium zur Erfassung und Bewertung der beruflichen Kompetenz von angehenden oder erfahrenen Lehrkräften nicht gibt (vgl. Terhart 2007), wird Lehrerarbeit schon immer folgenreich bewertet: Während und zum Abschluss des Referendariats werden angehende Lehrer von ihren Seminaarausbildern, z.T. auch von Ihren Ausbildungslehrern (Mentoren), in manchen Bundesländern auch durch die Schulleiter ihrer Ausbildungsschule bewertet. Wie sieht die Praxis der Bewertung aus? An welchen Kriterien und Standards orientieren sich die bewertenden Instanzen? Werden bestimmte Schwerpunkte gesetzt? Wird auf die Erfassung von *Kompetenzentwicklung* Wert gelegt? Gehen die eingesetzten Beurteilungsschemata auf individuelle Bemühungen einzelner Ausbilder zurück oder handelt es sich um gemeinsam entwickelte Kriterien? Wird in der Ausbildung für die verschiedenen Lehrämter (von der Grundschule bis zur Beruflichen Schule) unterschiedlich vorgegangen bzw. werden andere Akzente gesetzt?

Um einen ersten Einblick in die Praxis der Beurteilung von angehenden Lehrern zu bekommen, wurden im Sommer 2007 die 257 Studienseminare in Deutschland postalisch angeschrieben und gebeten, uns vorhandene Instrumente zugänglich zu machen, die zur Beurteilung von Referendaren eingesetzt werden.³ Die Adressen der Standorte wurden mithilfe einer Internetrecherche und in einigen Fällen durch eine Rückfrage bei der zuständigen Bildungsadministration des Bundeslandes erschlossen.⁴ Zusätzlich wurde ein entsprechender Aufruf in der Zeitschrift „SEMINAR. Lehrerbildung und Schule“, dem offiziellen Organ des Bundesverbandes der Seminar- und Fachleiter e.V., veröffentlicht. Die folgenden Beschreibungen und Analysen beruhen auf den Beurteilungsinstrumenten und Materialien, die uns von Fach- bzw. Studienseminarleitern und einigen Institutionen aus der Bildungsadministration (Ministerien oder Qualitätsagenturen der Länder) zugesandt wurden.

Insgesamt wurden uns 257 Dokumente zugesandt, wobei eine erste Durchsicht zeigte, dass ein Teil der Dokumente nicht direkt zur Beurteilung von Referendaren verwendet werden kann (z.B. allgemeine Programme der Studienseminare). Nachdem diese ausgeschlossen wurden, lagen noch 201 Beurteilungsinstrumente vor. Ergänzend stehen 87 ausgefüllte Informationsbögen zur Verfügung, in denen die Ausbilder gebeten wurden, einige Hintergrundinformationen zu den Instrumenten zu geben, beispielsweise

3 Die Versendung der Aufforderungen an die 40 Berliner Studienseminare erfolgte nicht postalisch, sondern via E-Mail durch die Bildungsadministration der Stadt.

4 Da die Ausbildung von Referendaren in Bayern nicht an Studienseminaren, sondern speziellen Ausbildungsschulen stattfindet, wurde für Bayern auf die „verbindlichen Anweisungen für die Beurteilung an Ausbildungsinstitutionen der zweiten Phase“, für das Lehramt an Grund-, Haupt-, Realschulen und Gymnasien zurückgegriffen, die uns freundlicherweise vom bayerischen Staatsministerium für Unterricht und Kultus zugänglich gemacht wurden.

die Schulform bzw. Lehramtsausbildung, in der die Instrumente eingesetzt werden oder Informationen zur Entwicklung der Instrumente.⁵

In Bezug auf die Regionen und Schulformen deckt das vorliegende Beurteilungsmaterial ein breites Feld ab. Mit wenigen Ausnahmen (Mecklenburg-Vorpommern, Saarland, Sachsen, Sachsen-Anhalt) liegen aus allen Bundesländern Instrumente vor. Auch die verschiedenen Lehrämter sind angemessen vertreten: Wir erhielten jeweils etwa 25 Beurteilungsschemata von Standorten, die auf das Lehramt an Beruflichen Schulen oder an Sonder- bzw. Förderschulen vorbereiten, etwa 50 Beurteilungsschemata aus der Ausbildung für das Lehramt an Gymnasien, ebenso viele aus der Ausbildung für das Lehramt an Grundschulen. Fast 100 Instrumente stammen aus Ausbildungsgängen, die auf das Lehramt an Haupt- und Realschulen vorbereiten.

Die gewählte Materialbasis führt zu gewissen Einschränkungen hinsichtlich der Aussagekraft der gewonnenen Ergebnisse: *Erstens* sei darauf verwiesen, dass aufgrund der vorliegenden Materialbasis keineswegs beansprucht werden kann, die Beurteilungsinstrumentarien im Rahmen der Zweiten Phase der Lehrerbildung repräsentativ für Deutschland abzubilden. Dies kann auch nicht Ziel einer ersten explorativen Studie zu dieser Fragestellung sein. *Zweitens* ist einschränkend anzumerken, dass die Analyse der eingesetzten Beurteilungsinstrumente zwar etwas über diese Instrumente aussagen und dadurch einen gewissen Einblick in die Beurteilungspraxis eröffnen. Über diese selbst könnten jedoch nur anhand ihrer *unmittelbaren* Beobachtung und Bewertung Aussagen getroffen werden.⁶ *Drittens* ist zu berücksichtigen, dass als Reaktion auf die Anschreiben bzw. den Aufruf uns vermutlich eher die elaborierteren Instrumente zugesandt worden sind und insofern in unserer Stichprobe überrepräsentiert sind.

Bei der Analyse der vorliegenden Materialbasis orientieren wir uns an einer Reihe von Dimensionen, die Porter u.a. (2001) für die Analyse von Beurteilungen von Lehrarbeit vorschlagen:

- Funktionen der Beurteilung (z.B. Ausbildungszugang, Bezahlung, Zertifizierung)
- Beurteilungsgegenstand (z.B. Fachwissen, Lehrfähigkeit, Lernzuwachs bei Schülern)
- Entwickler des Beurteilungsinstruments (z.B. Universitäten, Länder, Institute)
- Art der Beurteilung/Operationalisierung (z.B. Tests, Beobachtung, Portfolios)

5 Die ungleiche Zahl an Beurteilungsinstrumenten (201) und Informationsbögen (87) erklärt sich dadurch, dass oftmals in einer Zusendung mehrere Instrumente zusammengefasst und mit einem Informationsbogen versehen wurden, etwa wenn Seminarleiter zunächst intern die Instrumente ihrer Ausbilder zusammengetragen und dann an uns weiter geleitet haben.

6 Auch in der fallanalytischen Studie von Wernet (2008) zu den schriftlichen Abschlussgutachten der Ausbilder wird die unmittelbare Praxis des Beurteilens selbst nicht deutlich, sondern deren Resultat in Gestalt der Gutachten. Wernet vermisst eine kollegial-diskursive Orientierung in den Gutachten und erkennt stattdessen eine Orientierung der Gutachter an einem autoritativen, infantilisierenden oder technokratischen Konformismus.

3.2 Funktionen der Beurteilung: Zu welchem Zweck wird beurteilt?

An die überwiegend theoriebezogene universitäre Ausbildungsphase, die mit dem ersten Staatsexamen endet, schließt in allen Ländern der Bundesrepublik Deutschland die Zweite Phase an. Das Referendariat weist insofern zwei Lernorte auf: Studienseminar und Ausbildungsschule. In einer gewissen Weise ist diese zweite, berufspraktische Phase der Lehrerbildung, eine Art Berufseingangsphase (*induction phase*), d.h., sie ist einerseits eine begleitete Lernphase zum praktischen Berufseinstieg, andererseits dient sie aber auch der Prüfung, ob und in welchem Ausmaß sich die Referendare in Unterricht und Schule bewähren. Die Leistungen während des Referendariats gehen in die Bewertung am Ende dieser Phase, d.h. in die Note des zweiten Staatsexamens, ein. Die Noten aus dem ersten und zweiten Staatsexamen bilden dann die Basis für die Festlegung der Gesamtnote, die wiederum für die Einstellung eine zentrale Rolle spielt.

Welche Funktion erfüllen diese Beurteilungen während des Referendariats? Porter u.a. (2001, S. 259) unterscheiden hierzu zwischen Selektions-, Orientierungs- und Lern- bzw. Entwicklungsfunktion.

- Die Beurteilung hat eine *Selektionsfunktion*, wenn anhand von Bewertungsergebnissen darüber entschieden wird, ob ein Bewerber in ein Lehrerausbildungsprogramm aufgenommen wird, dieses fortsetzen kann oder von der Beurteilung der Übergang von der Ausbildung in den Beruf abhängt. Wenngleich das traditionelle Listenverfahren, das ausgebildete Lehrkräfte auf der Basis der Gesamtnote aus erstem und zweiten Staatsexamen, der jeweiligen Fächerkombination etc. in freie Lehrerstellen einweist, mittlerweile in allen Bundesländern, z.T. nur beschränkt auf bestimmte Schulformen, durch ein „schulscharfes“ Ausschreibungsverfahren abgelöst bzw. ergänzt worden ist (vgl. Hercher u.a. 2005; Schaefers/Terhart 2006), so hat sie doch immer noch eine wichtige Funktion: Die Bewerber um schulscharf ausgeschriebene Stellen werden zunächst immer noch wie bisher nach der Endnote bzw. Ordnungsgruppe sortiert, und unter den Besten kann die Schule dann nach den Anhörungen einen Besetzungsvorschlag machen.
- Als Zweites können Beurteilungskriterien einen *Orientierungsrahmen* aufspannen, in dem das Wissen und die Fähigkeiten benannt werden, die für die Ausübung des Lehrerberufs als wichtig und kritisch erachtet werden. So bieten transparente Beurteilungsmaßstäbe Ausbildern und Referendaren beispielsweise Orientierungshilfen zur Planung oder Durchführung von Unterricht.
- Eine *Personalentwicklungsfunktion* hat die Lehrerbeurteilung, sofern Kompetenzen innerhalb bestimmter Stufen der Ausbildung oder des Berufs weiter entwickelt werden sollen; insofern bildet Personalbeurteilung ein wichtiges Element innerhalb schulischer Personalentwicklung. Der kontinuierlichen Evaluierung der Wissens- und Fähigkeitsstände kommt im Rahmen neuerer Lernmodelle, in denen Lernen als selbstregulativer, konstruktivistischer Prozess verstanden wird, eine entscheidende Rolle zu (vgl. z.B. Bolhuis 2003).

In dem Informationsbogen wurde danach gefragt, welchen Zweck die Beurteilung aus Sicht der Ausbilder an den Studienseminaren hat. Zwei Drittel der Befragten (N=58) haben sich hierzu geäußert, die Antworten wurden analog zu den eben genannten drei Funktionsbereichen kodiert. Wenn beispielsweise angegeben wurde, dass die Instrumente im Examen, bei der Anfertigung von Zwischengutachten oder bei mündlichen oder schriftlichen Prüfungen eingesetzt werden, dienen sie der Selektion. Die Nachbesprechung von Unterrichtsbesuchen und weitere explizite Beratungs- und Rückmeldesituationen wurden der Personalentwicklungsfunktion zugerechnet. Eine Orientierungsfunktion haben solche Instrumente, wenn sie der Planungshilfe oder Zielformulierung dienen. Da z.T. mehr als eine Funktion angegeben wurde, waren Mehrfachkodierungen möglich bzw. notwendig.

Selektionsfunktion	Personalentwicklungsfunktion	Orientierungsfunktion	
60,3	62,1	12,1	Prozent*
35	36	7	Absolut*
*Mehrfachkodierung			

Abb. 2: Von den Ausbildern angegebene Funktion des Beurteilungsschemas (N=58)

Die Selektions- und die Personalentwicklungsfunktion erwiesen sich – aus Sicht der Ausbilder – als gleich bedeutsam: Jeweils circa 60 Prozent der Instrumente werden zu diesem Zweck eingesetzt, von diesen erfüllt die Hälfte beide Funktionen. Nur in 12 Prozent aller Fälle wird angegeben, dass die Instrumente zur Planung oder Zielformulierung eingesetzt werden, also der Orientierung der Referendare dienen. Dies ist erstaunlich, denn die Auszubildenden sollten eigentlich die konkreten Ausbildungsziele sowie die auf dem Wege dahin eingesetzten Beurteilungsraster kennen.⁷ Nicht umsonst ist das Setzen von Zielen und die zielgerichtete Planung ein entscheidender Aspekt innerhalb von Lernprozessen. Und da sich die Ausbildungsziele in den Beurteilungsinstrumenten niederschlagen (sollten), wäre es eigentlich geboten, diese den Lehramtsanwärtern transparent zu machen und sie so in ihrem Lernen zu unterstützen (vgl. hierzu auch Bolhuis 2003). Interessant ist in diesem Kontext, dass einige Instrumente explizit daraufhin ausgerichtet sind, die Entwicklungen der Referendare nachzuverfolgen. Einzelne Instrumente enthalten mehrere Beurteilungsfelder für die einzelnen Unterrichtsbesuche,

⁷ Die angeschriebenen Seminare reagierten unterschiedlich: In manchen Fällen erkundigten sie sich vorab bei der vorgesetzten Behörde und diese beim zuständigen Ministerium, ob man die Anfrage überhaupt beantworten dürfe. Andere Seminarleiter riefen an und drückten ihre Sorge aus, dass auf diesem Wege evtl. die Beurteilungskriterien den Referendaren bekannt werden könnten. Wieder andere Leiter von Studienseminaren meldeten sich und verwiesen auf die jedem zugängliche Homepage ihrer Einrichtung, von der man sich die (z.T. fachspezifischen) Beurteilungsschemata herunterladen könne.

anhand derer Entwicklungstrends unmittelbar sichtbar werden. In einem Fall machte die Leiterin eines Studienseminars für Berufskollegs ein ganzes Bündel von Beurteilungsinstrumenten zugänglich; in den Instrumenten zur Beurteilung einzelner Unterrichtsbesuche erweitern sich die Anforderungen (=Beurteilungsgegenstände) sukzessive.

3.3 Inhaltliche Bewertungsgegenstände: Was wird beurteilt?

Welche inhaltlichen Anforderungsbereiche werden durch die Beurteilungsschemata abgedeckt? Beurteilungskriterien können implizit ein Curriculum beschreiben, an dem sich Lehrende und Lernende orientieren, oder die Merkmale festlegen, anhand derer über die Eignung für den Lehrberuf entschieden wird. Insofern ist die inhaltliche Ausrichtung bzw. die Breite der abgedeckten Bereiche eine nicht unwichtige Determinante für die Qualität des Lehrberufs bzw. des Bildungssystems generell.

Um die Aspekte, die in den Beurteilungsinstrumenten bewertet werden, systematisch zu beschreiben, musste zunächst entschieden werden, ob aus dem vorliegenden Material heraus auf induktivem Weg ein Set von übergeordneten Kategorien gefunden werden sollte, das die genannten Inhalte im Wesentlichen beschreibt, oder ob anhand eines bereits vorliegenden Kategoriensystems bzw. eines theoretischen Rahmenmodells geprüft wird, inwieweit sich dessen Dimensionen im empirischen Material wieder finden lassen. Wir haben uns für die zweite Strategie entschieden. Die Kultusministerkonferenz hat mit den „Standards für die Lehrerbildung: Bildungswissenschaften“ einen verbindlichen Orientierungsrahmen geschaffen, in dem – bezogen auf das bildungswissenschaftliche Segment der Lehrerbildung⁸ – explizit für die beiden Phasen Kompetenzen für angehende Lehrerinnen und Lehrer benannt werden (vgl. KMK 2004). Diese Beschreibung der in der Ersten und der Zweiten Phase geforderten Kompetenzen dient uns als Grundlage für die inhaltliche Analyse der Beurteilungsinstrumente.

Den Standards liegt ein *umfassendes* Bild der Lehrerkompetenz zugrunde; innerhalb der Bereiche Unterrichten, Erziehen, Beurteilen und Innovieren werden elf Kompetenzen benannt, was angehende Lehrerinnen und Lehrer am Ende der ersten und zweiten Phase können sollten (vgl. Abb. 3).

Die inhaltliche Kategorisierung der einzelnen Beurteilungsinstrumente wurde auf Ebene der *elf Kompetenzen* durchgeführt. Jedes Dokument wurde darauf hin geprüft, welche der KMK-Standards sich dort wiederfinden. Während der Verkodung der Beurteilungsinstrumente bestätigte sich die Praktikabilität des Kodestammes. Allerdings erwies es sich als notwendig, diesen um die *vier Kompetenzbereiche* aus den KMK-Stan-

8 In den uns vorliegenden Beurteilungsschemata konnte kein Hinweis darauf gefunden werden, dass es neben den bildungswissenschaftlichen Standards spezifische, nur für Fachkompetenz oder fachdidaktische Kompetenz ausgerichtete Beurteilungskriterien gab. Die Unterteilung von Studiengangelementen (Fachstudien, Fachdidaktik, Erziehungswissenschaft, Praktika) ist für die universitäre Phase ein kennzeichnendes Element – in der Beurteilungspraxis der Zweiten Phase scheint diese Aufteilung hinfällig zu werden.

Kompetenzbereich	Kompetenz	Standard
1. Unterrichten	1. Unterricht planen, durchführen und reflektieren	etwa drei bis fünf Standards pro Kompetenz
	2. Lernsituationen gestalten	
	3. Selbstbestimmtes Lernen und Arbeiten fördern	
2. Erziehen	4. Lebensbedingungen kennen; individuell unterstützen	
	5. Normen, Werte und Eigenverantwortung vermitteln	
	6. Schwierigkeiten und Konflikte lösen	
3. Beurteilen	7. Diagnostizieren und Beraten	
	8. Transparente Beurteilung	
4. Innovieren	9. Bewusstsein des öffentlichen Amts	
	10. Lehrerberuf als ständige Lernaufgabe	
	11. Partizipation an schulischen Projekten	
12. Mitarbeit im Studienseminar (Teilnahme, Engagement und Verhalten im Studienseminar)		
13. Persönlichkeit (Höflichkeit, Pünktlichkeit, Humor, Äußeres, Ausstrahlung)		
14. Formale Kriterien zur Beurteilung von Hausaufgaben (Orthografie, Interpunktion, Schreibstil, Zitation)		

Abb. 3: KMK-Standards. Vier Bereiche und Kompetenzen

dards zu erweitern, da die Bewertungsgegenstände in einigen Instrumenten sehr abstrakt benannt wurden, sodass die differenzierte Zuweisung zu den einzelnen Kompetenzen nicht immer möglich war; wohl aber konnte eine Zuordnung zu den vier Kompetenzbereichen erfolgen. Und schließlich: Da entsprechende Hinweise in den uns zugesandten Beurteilungsschemata immer wieder auftauchten, wurde die Liste um die folgenden drei Anforderungsbereiche erweitert: Mitarbeit im Studienseminar, Persönlichkeit und formale Kriterien bei der Beurteilung der Hausarbeit. Auf diese Weise wurde für jedes Dokument geprüft, welche der 18 Aspekte darin beurteilt werden (4 Kompetenzbereiche, 11 Kompetenzen, 3 weitere Beurteilungsgegenstände).

Ein erster Eindruck hinsichtlich der Kriterien, nach denen angehende Lehrerinnen und Lehrer im Rahmen des Referendariats beurteilt werden, ergibt sich aus der Häufigkeit, in der die Kompetenzbereiche in den Beurteilungsinstrumenten vorkommen. Hierzu wurde für die vier *Bereiche* geprüft, ob sich wenigstens eine der zugehörigen Kompetenzen in den Instrumenten finden lässt.

Im Ergebnis wird deutlich, dass eine systematisch umfassende Beurteilung angehender Lehrerinnen und Lehrer eher selten ist: Lediglich 38 Prozent decken zwei oder mehr Kompetenzbereiche ab. Bei 20 Prozent der 201 vorliegenden Beurteilungsinstrumente

Typ I	Typ II	TYP III (38,3 %)			
Kein Bereich abgedeckt	Ein Bereich abgedeckt	Zwei Bereiche abgedeckt	Drei Bereiche abgedeckt	Vier Bereiche abgedeckt	
19,9 % 40	41,8 % 84	15,4 % 31	8,5 % 17	14,4 % 29	Prozent Absolut

Abb. 4: Umfang der abgedeckten Bereiche (N=201)

findet sich keiner und in den verbliebenen 42 Prozent nur einer der Bereiche wieder, die die KMK im Rahmen der Lehrerbildung für wichtig hält.

Hieran schließt sich die Frage an, was für ein Lehrerbild den Instrumenten zugrunde liegt. Wenn in den Beurteilungsschemata inhaltliche Beurteilungskriterien benannt werden, dann steht Unterrichten bzw. die Unterrichtskompetenz im Mittelpunkt fast aller dieser Instrumente. Somit bestätigt sich auch für den Bereich der Beurteilung von bzw. der Beurteilungsinstrumente für Referendaren *der Primat des Unterrichtens*, der auch aus anderen Forschungsprojekten zur Zweiten Phase der Lehrerbildung bekannt ist (vgl. Döbrich/Abs 2006, 2007, 2008; Frey 2008).

Fasst man diese ersten Eindrücke zusammen, so lassen sich aus der Perspektive der inhaltlichen Bewertungskriterien *drei Typen* ausmachen (vgl. Abb. 4):

- Beim *ersten Typus* wird überhaupt kein KMK-Kompetenzbereich und auch keine Einzelkompetenz benannt; 20 Prozent der Beurteilungsinstrumente sind diesem Typus zuzuordnen. Beispielsweise sind das Vordrucke, auf denen Ort und Zeit der Beurteilung sowie Daten zur Person notiert werden, für die eigentliche Beurteilung aber nur ein Freitextfeld vorgesehen ist.⁹ Entsprechende Instrumente erscheinen uns ungeeignet, wenn die Bewertung im Studienseminar systematisiert, fair und vergleichbar sein soll, da die Beurteilungskriterien, die ein Ausbilder heranzieht, letztendlich zufallsabhängig und unkontrolliert von Referendar zu Referendar wechseln können.
- Der *zweite Typus* ist demgegenüber einerseits als eine gewisse Verbesserung zu kennzeichnen, da immerhin für einen, und zwar *den zentralen Bereich der Lehrertätigkeit – das Unterrichten* – Beurteilungskriterien ausgearbeitet sind. Diesen zentralen Bereich zum *allein* differenziert erfassten bzw. beurteilten zu machen bedeutet jedoch andererseits, der Beurteilung ein eng geführtes, wenn nicht simplifiziertes Bild der Lehrertätigkeit zugrunde zu legen. So werden weitere ebenfalls bedeutsame Bereiche des Lehrerhandelns vernachlässigt; eine entsprechende Beurteilung bietet keine ausreichende Basis für evidenzbasierte Entscheidungen hinsichtlich der beruflichen Kompetenz der angehenden Lehrer. Dieser zweite Typus charakterisiert gut 40 Prozent der Instrumente.

⁹ Dass keine Kompetenzbereiche bzw. Einzelkompetenzen explizit ausgewiesen werden bedeutet nicht, dass diese bei den bewertenden Notizen, die ein Ausbilder in das Freifeld einträgt, nicht doch auftauchen können.

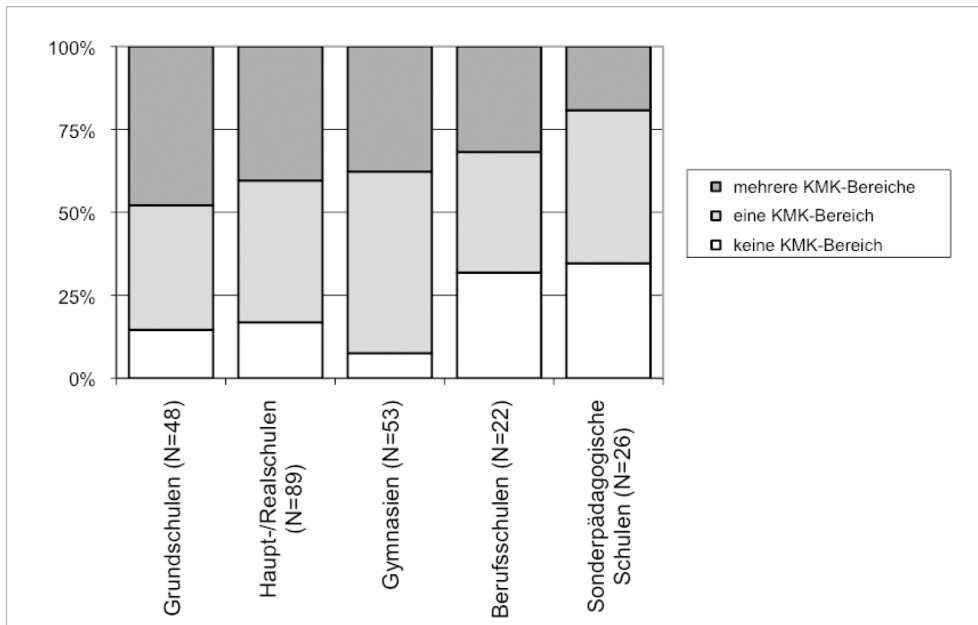


Abb. 5: Typen von Beurteilungsinstrumenten und Schulformbezug (N=238; Mehrfachzuordnungen notwendig)

- Erst der *dritte Typus* – und hierzu gehören nur 37 Prozent – ist durch eine umfassendere Beurteilung der beruflichen Kompetenz der Referendare gekennzeichnet, in der systematisch immerhin mehr (d.h. mehr als einer) Bereiche der Lehrerkompetenz erfasst werden. Sämtliche der vier von der KMK definierten Kompetenzbereiche werden lediglich von etwa 14 Prozent der uns vorliegenden Beurteilungsschemata abgedeckt. Dies ist sicherlich als bedenklich zu bewerten: Die weiter oben beschriebenen Funktionen vermag die Beurteilung der Referendare nur dann im Sinne des Lehrerbilds der KMK zu erfüllen, wenn angehende Lehrerinnen und Lehrer umfassend beurteilt werden, d.h. über die Kernaufgabe „Unterrichten“ hinaus.

Legt man die o.g. drei Typen von Beurteilungsinstrumenten zugrunde und fragt nach dem Schulformbezug, so ergibt sich folgendes Bild:

Demgemäß scheint mit „steigender“ Schulform der Anteil der ein breiteres Anforderungsprofil zugrunde legenden Beurteilungsschemata zurück zu gehen. Auffällig ist, dass in Sonder- bzw. Förderschulen dominierend keine spezifischen Kompetenzen oder nur ein Kompetenzbereich zugrunde gelegt werden. Hier sind allerdings weitere Analysen erforderlich.

Unterhalb der Ebene der vier Kompetenzbereiche erlaubt die Analyse der *einzelnen Kompetenzen* einen tieferen, differenzierteren Einblick in die Bewertungsgegenstände (vgl. Abb. 6). Bezugsbasis für die Prozentangaben in Abb. 6 ist die Gesamtzahl der Instrumente (N=201).

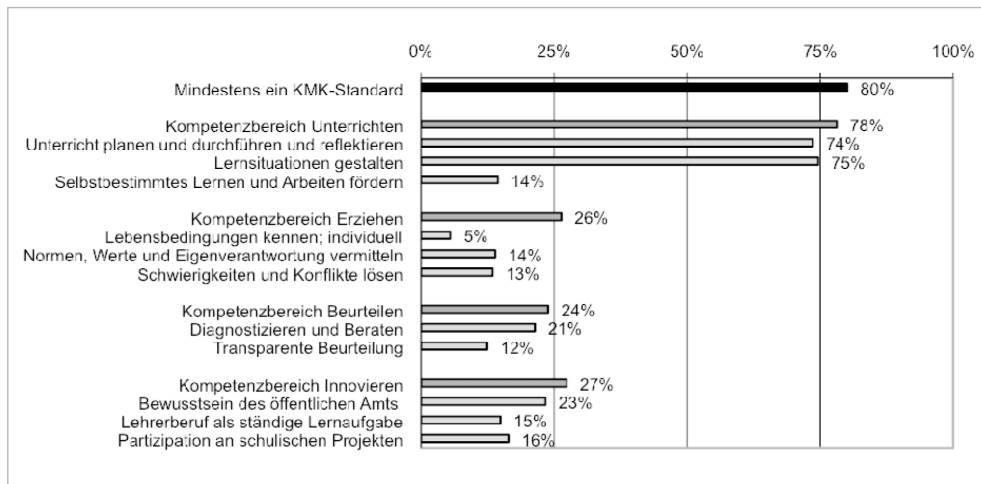


Abb. 6: Benannte Kompetenzbereiche bzw. Kompetenzen, Angaben in Prozent (N=201)

Wie bereits angeführt wurde, werden nur in 80 Prozent der Instrumente einer oder mehrere der KMK-Kompetenzbereiche benannt. Wenn Kompetenzbereiche benannt werden, dominiert der Bereich „*Unterricht*“; fast immer enthalten die Instrumente Fragen hierzu. Innerhalb dieses Bereichs jedoch zeigt sich, dass Referendare vergleichsweise selten dahin gehend eingeschätzt werden, wie es ihnen gelingt, *selbstbestimmtes Lernen* zu fördern. Offenbar legen die Ausbilder in der zweiten Phase weniger Wert darauf, selbstreguliertes Lernen in Unterricht und Schule zu fördern, als es aus Sicht der Bildungsforschung und -administration wünschenswert wäre.

Ebenso auffällig sind die Ergebnisse innerhalb des Bereichs „*Erziehen*“. So enthalten zwar ein Viertel der Instrumente Fragen, die diesem Bereich zuzuordnen sind, ein genauerer Blick zeigt allerdings, dass die *einzelnen* Kompetenzen viel seltener beurteilt werden. M.a.W., es wird anscheinend eher pauschal „die erzieherische“ Kompetenz bewertet. Ein bedenkliches Beispiel, das nebenbei auch die mangelnde Rezeption von Forschungsergebnissen manifestiert: Trotz der im Anschluss an die großen Schulleistungsstudien intensiv diskutierten herkunftsbedingten Disparitäten im deutschen Bildungswesen findet sich die vierte Kompetenz aus den KMK-Standards (2004) „Lehrerinnen und Lehrer kennen die sozialen und kulturellen Lebensbedingungen von Schülerinnen und Schülern und nehmen im Rahmen der Schule Einfluss auf deren individuelle Entwicklung“ seltener als alle anderen Kompetenzen wieder. In nur 5 Prozent der Fälle ist die Beurteilung von Referendaren sensibel für diesen Bereich!

Interessant ist auch eine nähere Betrachtung der Kompetenzen im Bereich „*Beurteilen*“. Die Diagnose individueller Leistungsstände gehört zum Alltagsgeschäft von Lehrkräften und erfüllt vielerlei Funktionen; so dient sie der Notengebung und ist bzw. sollte selbst Teil des Lernprozesses sein (z.B. Monitoring über die Leistungsstände, Grundlage für adaptives Unterrichten). Vor diesem Hintergrund ist es erstaunlich, dass nur jedes vierte Instrument diesen Bereich zumindest in Teilen abdeckt. Besonders kritisch ist

aber, dass nur halb so oft (12 %) in den Beurteilungsschemata thematisiert wird, wie es um die Fähigkeiten eines angehenden Lehrers hinsichtlich der Beurteilung von Schülerleistungen bestellt ist. Dabei ist die schon beinahe sprichwörtliche „Fragwürdigkeit der Zensurengebung“ (K. Ingenkamp) seit Jahrzehnten in Wissenschaft und Schulpraxis bekannt. In diesem Kontext ist auf die seit den großen Leistungsvergleichsstudien bemängelte unzureichende diagnostische Kompetenz der Lehrkräfte hinzuweisen. Das Berücksichtigen von Lernausgangslagen, das Erkennen und Zuordnen von Lernschwierigkeiten, die sach- und individuengerechte Beurteilung von Schülerleistungen etc. – dieses Aufgabenfeld eines jeden Lehrers wird nur von jedem zehnten Beurteilungsschema erfasst!

Der Kompetenzbereich „*Innovieren*“ setzt sich aus dem Bewusstsein des öffentlichen Amtes, der individuellen und organisationalen Weiterentwicklung zusammen. In den Beurteilungsinstrumenten liegt der Schwerpunkt auf dem erstgenannten Punkt, hierzu zählt unter anderem, wie Referendare mit Belastung umgehen, wie sie ihre Arbeit (individuell und im Kollegium) organisieren, ob sie mit dem Schulrecht vertraut sind. Seltener wird beurteilt, ob ein Referendar den Lehrerberuf als ständige Lernaufgabe auffasst, ob er sich an der Schulentwicklung aktiv beteiligt und an schulischen Projekten partizipiert.

Zusammenfassend kann festgehalten werden, dass die ersten Deskriptionen auf der *Ebene der vier Kompetenzbereiche* die Unterscheidung von drei Typen von Instrumenten erbracht hat. Ein Blick auf die einzelnen Kompetenzen *unterhalb der Ebene der Kompetenzbereiche* zeigt, dass diese durchaus unterschiedlich häufig bei der Beurteilung zugrunde gelegt werden. So wird im Bereich Unterrichten die durchaus anspruchsvolle Kompetenz, selbstbestimmtes Lernen fördern zu können, eher vernachlässigt. Alte und neue einschlägige Forschungsbefunde der Schul- und Unterrichtsforschung lassen sich in den uns vorliegenden Beurteilungsschemata oftmals nicht wieder finden. So wird z.B. selten beurteilt, ob Referendare die Lebensbedingungen ihrer Schülerinnen und Schüler kennen und berücksichtigen, ob sie Beurteilungskriterien transparent machen können, inwiefern sie sich an Schulentwicklung beteiligen etc. Die KMK-Kompetenzbereiche und Standards – hier genutzt als theoretisches Rahmenmodell für die inhaltliche Analyse und zugleich ein verbindlicher Rahmen für die Ausbilder an den Studienseminaren – werden nur in 14 Prozent der Beurteilungsinstrumente vollständig abgedeckt; anders formuliert: Den allermeisten Instrumenten liegt ein in Richtung auf „Unterricht“ eng geführtes, nicht die Breite der beruflichen Anforderungen insgesamt abdeckendes Beurteilungskonzept zugrunde.

Wie bereits erwähnt, wurde angesichts des zugesandten Materials der Kodestamm um die Punkte „Mitarbeit im Studienseminar“, „Persönlichkeit“ und „formale Kriterien bei der Beurteilung der Hausarbeit“ erweitert. Formale Kriterien bei der Beurteilung der Hausarbeit finden sich naturgemäß in Instrumenten, die speziell zur Bewertung schriftlicher Arbeiten herangezogen werden. In insgesamt zwölf Beurteilungsinstrumenten werden die Orthografie, Interpunktion, Schreibstil oder Korrektheit der Zitation bewertet. Die Mitarbeit im Studienseminar fließt bei 13 Prozent der Instrumente in die Bewertung ein und kommt damit mehr als doppelt so häufig vor wie etwa die KMK-Kompe-

Zusammenfassung der Nachbesprechung im Fach Klasse/Kurs: Ausbildungslehrer(in):		am: Zeit	Name der Referendarin/ Name des Referendars	Nr. des UB
	beibehalten/ Routinen bilden	weiterent- wickeln/ verändern	besondere Aspekte	nächster UB
PLANUNG				
DURCHFÜHRUNG: z.B. L - S Interaktion Lernzuwachs				
Reflexion				

Abb. 7a: Beispiel A (Instrument eines Fachleiters für Biologie an Gymnasien/Gesamtschulen)

tenz „Lebensbedingungen kennen“. Seltener wird die Persönlichkeit der angehenden Lehrerinnen und Lehrer beurteilt: in 14 von 165 Instrumenten werden Merkmale wie Höflichkeit, Pünktlichkeit, Humor, Äußeres, Ausstrahlung etc. in die Beurteilung mit- einbezogen. Obwohl nicht von der Hand zu weisen ist, dass auch solche Faktoren für die spätere Berufssituation eines angehenden Lehrers – positiv wie negativ – von Bedeu- tung sein können, kann man mit guten Gründen vermuten, dass es sehr schwierig ist, hierfür auch nur annähernd zuverlässige Beurteilungsprozeduren zu finden. Insofern kann man die geringe Zahl der Beurteilungsinstrumente, in denen diese Persönlichkeits- merkmale beurteilt werden, durchaus kontrovers beurteilen.

Kompetenz 4: Lehrerinnen und Lehrer reflektieren ihren Unterricht und tragen damit zur kontinuierlichen Optimierung ihrer Unterrichtsplanung und -durchführung bei. <i>Die Lehrerin/der Lehrer...</i>	
Standard	Kompetenznachweis
vergleicht die angestrebten Lernziele mit dem erreichten Lernzuwachs.	
bewertet die didaktischen und methodischen Planungsentscheidungen im Hinblick auf ihre Wirksamkeit.	
zeigt gegebenenfalls realistische Handlungsalternativen auf.	
begründet und bewertet Planungsabweichungen	
schätzt die Qualität der Unterrichtsorganisation und -moderation ein.	
beurteilt die unterrichtliche Kommunikation und Interaktion.	
nutzt die Ergebnisse der eigenen und der gemeinsamen Reflexion für die Verbesserung des Unterrichts.	
Kompetenzbereich 2: Erziehen Lehrerinnen und Lehrer üben ihren Erziehungsauftrag aus. Kompetenz 5: Lehrerinnen und Lehrer kennen die sozialen und kulturellen Lebensbedingungen von Schülerinnen und Schülern und nehmen Einfluss auf deren Entwicklung <i>Die Lehrerin/der Lehrer...</i>	
Standard	Kompetenznachweis
erfasst etwaige Benachteiligungen von Schülerinnen und Schülern beim Lernprozess und ergreift individuelle Maßnahmen der pädagogischen Hilfe und Prävention.	
beachtet kulturelle Unterschiede sowie die Bedeutung geschlechtsspezifischer Einflüsse bei der Gestaltung von Bildungs- und Erziehungsprozessen in der jeweiligen Lerngruppe.	
fordert von den Schülerinnen und Schülern Leistungen im Rahmen ihrer individuellen Fähigkeiten.	
Kompetenz 6: Lehrerinnen und Lehrer vermitteln Werte und Normen und unterstützen selbstbestimmtes Urteilen und Handeln von Schülerinnen und Schülern. <i>Die Lehrerin/der Lehrer...</i>	
Standard	Kompetenznachweis
vermittelt die Wertvorstellungen, die dem Grundgesetz, der niedersächsischen Verfassung und dem Bildungsauftrag des NSchG zugrunde liegen.	
vereinbart mit den Schülerinnen und Schülern Regeln für das Verhalten im Unterricht und in der Schule und achtet auf ihre Einhaltung.	
fördert bei den Schülerinnen und Schülern eigenverantwortliches Urteilen und Handeln auf der Grundlage wertbewusster Handlungen.	
leitet die Schülerinnen und Schüler zu einem kritischen und verantwortungsvollen Umgang mit Medien an.	

Abb. 7b: Beispiel B (Ausschnitt aus dem Instrument der Fachleiter eines Studienseminars für Grund-, Haupt- und Realschulen)

3.4 Art der Beurteilung: Wie wird beurteilt?

Angehende Lehrerinnen und Lehrer können während und am Ende des Referendariats auf unterschiedliche Art und Weise begutachtet werden. In Deutschland sind Lehrproben, mündliche Prüfungen und schriftliche Hausarbeiten (z.B. Unterrichtsentwürfe) die dominierenden Beurteilungssituationen. Porter u.a. (2001) fassen solche Beurteilungssituationen unter dem Oberbegriff *Beobachtung und Befragung* zusammen, zudem unterscheiden sie zwischen standardisierten Tests (Multiple-Choice und offene Fragen), Portfolios und Assessment-Centern als weitere Formen zur Lehrerbeurteilung.

Anders als in den Vereinigten Staaten¹⁰ werden standardisierte Tests wie der Pre-Professional Skills Test (PPST) (vgl. Mitchell u.a. 2001, Kap. 5) im Rahmen der deutschen Lehrerausbildung nicht eingesetzt. So befindet sich auch unter den uns vorliegenden Materialien kein Test oder auch nur ein testähnliches Instrument. Ebenso wurden uns keine Instrumente oder Verfahrensbeschreibungen zugesandt, die nahe legen würden, dass die Erfassung von Lehrerkompetenzen in Form eines Assessment-Centers durchgeführt wird. Eine neuere Form, die Arbeit von Lehrerinnen und Lehrern zu dokumentieren, sind Portfolios (vgl. Bird 1990; Wolf u.a. 1997). Sie bestehen aus Dokumenten wie Videomitschnitten, Dokumentationen der Unterrichtsplanung, Beispielstunden, Bewertungen, Kommentaren, Beispielen von betreuter Schülerarbeit etc. und sollen eine individuelle, authentische Beurteilung ermöglichen. Nur ein einziges der uns vorliegenden Dokumente greift die Idee der Portfolios auf: An einem Studienseminar für das Lehramt an Grund- und Hauptschulen wurde kooperativ und sorgsam ein *Handbuch zur Portfolioarbeit* entwickelt, in dem beschrieben wird, wie Referendare dies anlegen und für die eigene Weiterentwicklung nutzen können. Hier erfüllt das Instrument klar eine Lern- und Entwicklungsfunktion: angehende Lehrerinnen und Lehrer sollen sich ihre eigenen Kompetenzen bewusst machen und Entwicklungen nachvollziehen.

Wie bereits angesprochen, sind Beobachtungen und Befragungen die Regel, wenn angehende Lehrer in Studienseminaren beurteilt werden. Nachfolgend sollen die Beurteilungsinstrumente, die in den Studienseminaren hierzu entwickelt wurden und eingesetzt werden, in Bezug auf die Art und Operationalisierung etwas ausführlicher beschrieben und diskutiert werden. So können einzelne inhaltliche Bewertungsbereiche exakt und detailliert oder aber sehr abstrakt beschrieben werden. In Freitextfeldern können Prüfer individuelle Beurteilungstexte anfertigen, wohingegen die quantitative Einstufung auf einer Antwortskala interindividuelle Vergleiche erleichtert.

Ein Beispiel für einen Bewertungsbogen ist der eines Fachleiters für Biologie an Gymnasien/Gesamtschulen (Abb. 7, Beispiel A), es handelt sich um einen Vordruck, auf dem, neben Feldern für Name, Klasse, Datum etc., zwischen den Bereichen Planung, Durchführung und Reflexion unterschieden wird. Was etwa unter Durchführung verstanden wird, bleibt unklar. So könnten Beurteilende, Beurteilte und Dritte (etwa Schulleiter) unter „Durchführung“, „Didaktik“ und/oder „Unterrichtsmethoden“ und/oder

¹⁰ Einen Überblick über den Einsatz und die Problematik von Tests findet man beispielsweise bei Peterson (2000) oder Porter u.a. (2001).

„Medieneinsatz“ usw. jeweils etwas anderes verstehen. Jedenfalls liegt es beim Ausbilder, was er in die Freitextfelder einträgt und welche Qualitäten er beurteilt.

Sehr viel genauer definieren die Fachleiterinnen und -leiter eines Studienseminars für Grund-, Haupt- und Realschulen ihre Bewertungsgegenstände und orientieren sich dabei inhaltlich an den von der KMK veröffentlichten Standards für die Lehrerbildung (vgl. Abb. 7, Beispiel B). Angehende Lehrerinnen und Lehrer werden in den Kompetenzbereichen Unterrichten, Erziehen, Beurteilen und Innovieren beurteilt, wobei sich diese Bereiche in jeweils zwei bis vier Kompetenzen aufgliedern, die ihrerseits wiederum anhand von jeweils drei bis elf Indikatoren definiert werden. Beispielsweise wird im Bereich Unterrichten die Kompetenz fach- und sachgerechte Unterrichtsplanung anhand von Indikatoren wie Berücksichtigung curricularer Vorgaben, Dokumentation der kurz-, mittel- und langfristigen Unterrichtsplanung oder gezielte Methodenwahl bewertet. Für jeden einzelnen Indikator ist ein Freitextfeld vorgesehen, in das der „Kompetenznachweis“ eingetragen wird.

3.5 Formen der Beurteilung

Freitextkommentare: Bei beiden Beispielen (vgl. Abb. 7) verfassen die Beurteiler Freitextkommentare. Zu mehr oder weniger eng spezifizierten Bereichen beurteilen Ausbilder angehende Lehrerinnen und Lehrer mit individuell angefertigten Texten bzw. Kommentaren. Ein Vorteil dieser Art der Bewertung ist, dass sie genügend Freiraum lässt, um individuelle Stärken und Schwächen von Referendaren festzuhalten. Die Spezifika einzelner Lehramtsanwärter und besondere Entwicklungsschwerpunkte können konkret beschrieben werden.

Ratingskalen: Alternativ dazu können die angehenden Lehrerinnen und Lehrer auch auf einer Ratingskala eingeschätzt werden. Dabei wird im Itemstamm der Bewertungsgegenstand benannt, und anhand einer Antwortskala urteilt der Ausbilder ob (dichotom) oder in welchem Maß (abgestuft) die Lehramtsanwärterinnen und -anwärter bestimmte Verhaltensweisen oder Fähigkeiten zeigen. Während Freitextfelder individuell angepasste Beurteilungen ermöglichen, eröffnen Ratingskalen den Vergleich unterschiedlicher Referendare, und die einzelnen Beurteilungen können (ggf. gewichtet) zusammengefasst werden.

Während ein Ausbilder anmerkt: „Die zur Unterrichtsbeurteilung unerlässlichen Gewichtungsprozesse können in diesem Referenzrahmen nicht dargestellt werden“, zeigt sich im Bewertungsmanual, welches der Leiter eines Studienseminars für das Lehramt an Realschulen zur Abschlussbegutachtung im Rahmen der zweiten Staatsprüfung einsetzt, dass die Beurteilung inklusive der darin enthaltenen Gewichtungsprozesse durchaus systematisiert werden können: Fünf Bereiche werden anhand von jeweils 4 bis 15 Indikatoren bewertet. Anhand der Indikatoren werden die Teilbereiche zunächst einzeln benotet, diese werden dann gewichtet und zu einer Gesamtnote zusammengefasst. Beispielsweise wird das dienstliche Verhalten dabei einfach und das fachdidaktisch-pädagogische Wissen dreifach gewertet.

Kompetenzniveaus: Die detaillierte Beschreibung von Kompetenzniveaus bietet eine seltenere, nichtsdestoweniger interessante Form der Begutachtung. „Es geht hierbei z.B. um die Frage, welche spezifischen Anforderungen eine Person mit einer hohen Kompetenz bewältigen kann und welche Anforderungen eine Person mit einer niedrigen Kompetenz gerade noch bewältigt und welche nicht“ (Hartig/Klieme 2007, S. 133). So werden in einem fachübergreifend eingesetzten Beurteilungsinstrument für angehende Gymnasiallehrerinnen und -lehrer jeweils drei Niveaustufen für 22 Kompetenzen beschrieben. Im Bereich Didaktik wird bewertet, wie Lernprozesse transparent gemacht werden. Das untere Kompetenzniveau zeichnet sich etwa durch isolierte unterrichtliche Schritte, fehlende Zusammenhänge oder Kontextuierungen aus, der Unterricht von Referendaren des zweiten Niveaus ist dahingegen sinnvoll aufeinander aufbauend, wobei dies den Schülern allerdings nicht oder erst spät klar wird. Dies gelingt Referendaren auf dem höchsten Kompetenzniveau: Einzelne Unterrichtsschritte fügen sie zusammen und zudem wird von Beginn an nachvollziehbar dargestellt, wie sich die einzelnen Stunden in den Unterrichtskontext einbetten. Die Verortung auf einem Niveau ist sehr anschaulich und wohl auch aussagekräftiger als die auf einer Skala von „trifft zu – trifft eher zu – trifft eher nicht zu – trifft nicht zu“.

Gleichzeitig scheint es in der Praxis schwierig zu sein, geeignete Kompetenzstufen zu finden, die eine Kompetenz treffend charakterisieren und der Praxis standhalten. Wenngleich das in Abb. 9 angeführte Beispiel von Kompetenzniveaus zu den gelungenen gezählt werden darf, mag es auch hier schwierig sein, einen Referendar exakt auf einer Kompetenzniveau zu verorten, der zwar häufig Überblicke gibt, dessen Stunden sich für die Schülerinnen und Schüler dennoch nicht sinnvoll zusammenfügen. Gleichzeitig besteht die Gefahr, dass bei dem anspruchsvollen Vorhaben, Kompetenzniveaus zu formulieren, de facto Ratingskalen konstruiert werden, etwa wenn sich die unterschiedlichen Niveaus nur durch Begriffe wie „regelmäßig – punktuell – in Ausnahmefällen“ unterscheiden, und nicht wirklich unterschiedliche Qualitäten durch die Niveaus bzw. deren Beschreibung ausgedrückt werden.

Quantitative/qualitative Beurteilung: Für die vorliegenden Beurteilungsinstrumente wurde systematisch untersucht, ob Referendare qualitativ (Freitextkommentare) oder quantitativ (Ratingskalen und Kompetenzniveaus) beurteilt werden. Es erwies sich als notwendig, dieser Unterscheidung zwei weitere Kategorien hinzuzufügen: Hybridformen, in denen qualitative und quantitative Bewertungsformen kombiniert werden, sowie eine Kategorie für Instrumente, bei denen unklar bleibt, wie die Beurteilungsmanuale letztendlich genutzt werden, da keinerlei freie Felder für die Beurteilung vorgesehen sind. Dabei handelt es sich um aneinander gereihte, teilweise inhaltlich strukturierte Stichpunkte. Ob und wie diese Aspekte in der Beurteilungspraxis bewertet werden bleibt unklar; es kann vermutet werden, dass diese Listen wenig verbindlich bzw. beliebig herangezogen werden.

Im Ergebnis werden Referendare bei 15 Prozent der uns zugesandten Beurteilungsinstrumente qualitativ beurteilt, etwa doppelt so groß ist der Anteil der quantitativen Formen der Bewertung, wobei Ratingskalen drei mal häufiger eingesetzt werden als Kompetenzstufen. Sehr selten sind hingegen (4 Prozent) Hybridformen, in denen quali-

Kompetenz	Qualitätsstufe 1	Qualitätsstufe 2	Qualitätsstufe 3
Freiraum für Schüleraktivitäten wird gewährt.	Die Schüler erhalten durchgängig Gelegenheit zu zielgerichteter Aktivität und Eigenverantwortung. Das selbstständige Lernen wird gefördert.	Die Schüler erhalten für einzelne Phasen Gelegenheit zu zielgerichteter Aktivität und Eigenverantwortung. Möglichkeiten des selbstständigen Lernens unter Berücksichtigung innovativer Unterrichtsverfahren werden punktuell eröffnet.	Die Schüler erhalten nur durch einen hohen Anteil an Lehrersteuerung Gelegenheit zu eigener Aktivität. Möglichkeiten des selbstständigen Lernens werden nur punktuell eröffnet.
Das selbstständige Lernen wird gefördert.	Das selbstständige Lernen wird im Rahmen einer Unterrichtssequenz regelmäßig gefördert.	Möglichkeiten des selbstständigen Lernens werden punktuell gefördert.	Möglichkeiten des selbstständigen Lernens werden nur in Ausnahmefällen gefördert.
Das Lehrerverhalten ist von pädagogischer Zuwendung geprägt.	Das Lehrerverhalten ist konsequent und den einzelnen Schülerinnen und Schülern sowie der gesamten Lerngruppe zugewandt. Der Lehrer schafft durch Freundlichkeit und Bestimmtheit eine angenehme und effektive Lernatmosphäre, die von Empathie und Hilfsbereitschaft geprägt ist.	Das Lehrerverhalten ist freundlich und den einzelnen Schülerinnen und Schülern sowie der Lerngruppe insgesamt zugewandt. In ihrem Auftreten ist die Lehrkraft gelegentlich aber nicht konsequent genug, so dass sich Unsicherheiten ergeben.	Das Lehrerverhalten ist nicht angemessen distanziert. Das Verhalten der Lehrkraft ist häufig inkonsequent bzw. nicht situationsangemessen. Der Lehrer kann sich bei Disziplinproblemen nicht genügend durchsetzen. Er reagiert zum Teil unfreundlich auf Hinweise der Schüler.

Abb. 8: Drei ausgewählte Beispiele für Kompetenzniveaus; entnommen aus einem Instrument, das alle Ausbilder eines Studienseminars für das Lehramt an Gymnasien einsetzen.

tative und quantitative Formen der Bewertung kombiniert werden. Die Gruppe derjenigen Instrumente schließlich, bei denen nicht spezifiziert wird, wie sich die inhaltlichen Kriterien in der Beurteilung nieder schlagen, ist die mit Abstand größte: *In jedem zweiten Instrument bleibt unklar, wie die Referendare letztendlich bewertet werden.* Es kann vermutet werden, dass die Beurteilungskriterien unverbindlich sind und nicht systematisch zur Beurteilung herangezogen werden.

Aus der Art der Operationalisierung ergeben sich bedeutsame Implikationen für die Brauchbarkeit von Beurteilungsinstrumenten. Werden die Kompetenzen und Fähigkeiten der Referendare anhand von Freitextfeldern beurteilt, so können die Ausbilder individuelle Beurteilungen anfertigen. Insbesondere dann, wenn individuelle Entwicklungs-

Qualitativ	Quantitativ	Hybrid	Keine	
14,9 % 30	28,9 % 58	3,5 % 7	52,7 % 106	Prozent Absolut

Abb. 9: Beurteilungsform (N=201)

schwerpunkte vereinbart werden, hat diese Art der Beurteilung klare Vorzüge. Sollen Referendare hingegen miteinander verglichen werden, so ist dies anhand von Freitextkommentaren schwierig; Ratingskalen oder Kompetenzstufen haben demgegenüber den Vorteil, dass der Beurteilungsrahmen (Abstufungen) exakt vorgegeben ist, wodurch die Beurteilungen direkt miteinander vergleichbar werden. Die Kombination qualitativer und quantitativer Formen der Beurteilung verknüpft die Vorteile beider Formen; allerdings wird das in den uns vorliegenden Instrumenten nur sehr selten praktiziert.

4. Schlussbemerkung

Auf der Basis des vorliegenden Materials und unter Berücksichtigung der bisher durchgeführten Analysen wird deutlich, dass in der Zweiten Phase der Lehrerbildung bei der Beurteilung angehender Lehrerinnen und Lehrer durch die Ausbilder durchaus Kriterienlisten, Beurteilungsmanuale, mehr oder weniger differenzierte Evaluationsbögen etc. eingesetzt werden. Der Charakter und die Qualität der uns zur Verfügung gestellten Beurteilungsschemata variierten sehr stark. Eine konkrete Operationalisierung des Beurteilungsaktes selbst ist nur in sehr wenigen Instrumenten zu erkennen: bei der Hälfte fanden sich dazu überhaupt keine Angaben. In zwei Drittel der Fälle scheinen diese Kriterien gemeinsam (innerhalb einer Fachgruppe, innerhalb eines Studienseminars) erarbeitet zu sein. Dies ist erfreulich, macht aber andererseits auch deutlich, dass ein Drittel anscheinend jeweils von nur einem (!) Ausbilder (zur Organisation und Dokumentation seiner eigenen Beurteilungsarbeit) entwickelt worden ist. Dies macht die Klage vieler Referendare über uneinheitliche Beurteilungskriterien und -standards verständlich.

Dass die Arbeit im Referendariat und auch der Beurteilungsschwerpunkt der Beurteiler sehr eindeutig auf die Entwicklung der didaktischen Kompetenz, der Fähigkeit zur routinisierten „Herstellung“ von Unterricht gerichtet ist, ist mit Blick auf die Breite des Aufgabenspektrums von Lehrern sicherlich eine Engführung, sollte jedoch auch nicht überraschen: Die Unterrichtsaufgabe steht zu Recht im Mittelpunkt der Lehrertätigkeit, die zuverlässige „Erzeugung“ von Unterricht ist gewissermaßen das funktionale Minimum, das jede Lehrkraft nach Übernahme in den Dienst beherrschen muss. Die weiteren und z.T. anspruchsvolleren inner- und ausserunterrichtlichen Aufgaben des Lehrerberufs sollten durch die Ausbildung sicherlich vorbereitet und angebahnt sein; ihre Beherrschung kann jedoch bei Berufsanfängern noch nicht vorausgesetzt werden. Auch die beste Lehrerbildung kann nicht den „kompletten“ Lehrer erzeugen, sondern sollte darauf gerichtet sein, möglichst gut vorbereitete Berufsanfänger in die Berufspraxis zu entlassen, die dort weiter lernfähig sind – und bleiben. Insofern wird durch die Beurteilungspraxis während des Referendariats noch einmal nahegelegt, dass dem lebenslangen Weiterlernen auch im Lehrerberuf ein hoher Stellenwert zukommt.

Die Analyse macht deutlich, dass die gegenwärtig praktizierte Beurteilung von Referendaren zwar im Resultat zu den notwendigen Beurteilungsergebnissen führt: Zwischen- und Examensnoten kommen zustande; das funktionale Minimum wird erfüllt. Aus kompetenzdiagnostischer Sicht muss die Qualität dieses Prozesses selbst jedoch

zumindest als problematisch, wenn nicht in manchen Varianten als unzureichend beurteilt werden. Dies ist *nicht* als moralisierender Vorwurf an die Akteure, sondern als dringende Aufforderung an die für das System verantwortlichen Instanzen zu verstehen, die bei der Lizenzierung von zukünftigen Lehrkräften eingesetzten formativen und summativen Beurteilungsinstrumente und Evaluationsprozeduren dringend weiter zu entwickeln. Die Bedeutsamkeit des Faktors „Lehrer“ für die Qualität von Unterricht und Schule macht dies unumgänglich.

Literatur

- Allemann-Ghionda, C./Terhart, E. (Hrsg.) (2006): Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehren: Ausbildung und Beruf. (51. Beiheft der Zeitschrift für Pädagogik). Weinheim: Beltz.
- Bird, T. (1990): The Schoolteachers' Portfolio: An Essay on Possibilities. In: Millman, J./Darling-Hammond, L. (Hrsg.): The New Handbook of Teachers Evaluation: Assessing Elementary and Secondary School Teachers. Newbury Park: Sage, S. 241–256.
- Blömeke, S./Kaiser, G./Lehmann, R. (2008): Professionelle Kompetenz in Lehrerbildung und Lehrerberuf. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung. Münster: Waxmann.
- Blömeke, S./Reinhold, P./Tulodziecki, G./Wildt, J. (Hrsg.) (2004): Handbuch Lehrerbildung. Bad Heilbrunn: Klinkhardt.
- Böhner, M. (2008): Wie wirkt sich der Vorbereitungsdienst auf die professionelle Kompetenz von Lehramtsanwärter/innen aus? Ein Blick in die Forschungslage. In: SEMINAR 14, H. 3, S. 81–95.
- Bölting, F.-J./Thomas, S. (2007): Das Referendariat als Ausbildungsphase der Berufseinführung. In: Óhidy, A./Terhart, E./Zsolnai, J. (Hrsg.): Lehrerbild und Lehrerbildung. Praxis und Perspektiven der Lehrerausbildung in Deutschland und Ungarn. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 203–217.
- Bolhuis, S. (2003): Towards Process-Oriented Teaching for Self-Directed Lifelong Learning: A Multidimensional Perspective. In: Learning and Instruction 13, S. 327–347.
- Döbrich, P./Abs, H.J. (2006): Pädagogische Entwicklungsbilanzen mit Studienseminaren in Hessen (PEB-Sem). In: SEMINAR 1, S. 93–100.
- Döbrich, P./Abs, H.J. (2007): Bewertungen in der zweiten Phase der Lehrerbildung – aktuelle Ergebnisse der Pädagogischen Entwicklungsbilanzen mit Studienseminaren in Hessen. In: SEMINAR, S. 20–31.
- Döbrich, P./Abs, H.J. (2008): Evaluation der zweiten Phase der Lehrerbildung. Pädagogische Entwicklungsbilanzen mit Studienseminaren in Hessen. Frankfurt: DIPF.
- Englert, R./Porzelt, B./Reese, A./Stams, E. (2006): Innenansichten des Referendariats. Wie erleben angehende Religionslehrerinnen an Grundschulen ihren Vorbereitungsdienst? Münster: LIT.
- Frey, A. (2008): Kompetenzstrukturen von Studierenden in der ersten und zweiten Phase der Lehrerbildung. Eine nationale und internationale Standortbestimmung. Landau: Verlag Empirische Pädagogik.
- Hartig, J./Klieme, E. (2007). Kompetenz und Kompetenzmessung. In: Schweizer, K. (Hrsg.): Leistung und Leistungsdiagnostik. Berlin: Springer, S. 127–143.
- Hercher, J./Schaefer, C./Treptow, E./Rothland, M. (2005). „Jeder Schule ihre Lehrer?“. Empirische Befunde zum Auswahlverfahren bei der Besetzung von Lehrerstellen in Nordrhein-Westfalen. In: Zeitschrift für Erziehungswissenschaft 8, S. 305–320.

- Hilligus, A./Rinkens, D. (Hrsg.) (2006): Standards und Kompetenzen – neue Qualität in der Lehrerausbildung? Münster: LIT-Verlag.
- Hoppenworth, U. (1993): Der Unterrichtsbesuch. Implizite Unterrichtstheorien von Ausbildern und Auszubildenden in der Zweiten Phase der Lehrerausbildung. Frankfurt a.M.: Lang.
- Katzenbach, D. (1999): „Die schlimmste Zeit meines Lebens“. Das Leiden am Referendariat (Wie) kann Supervision helfen. In: Pädagogik 10, S. 29–32.
- KMK (2004): Standards für die Lehrerbildung. Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004. Bonn: KMK; u.a. abgedruckt in: Zeitschrift für Pädagogik 51, 2005, S. 280–290.
- Lehmermöhle, D./Jareis, D. (Hrsg.) (2003): Professionalisierung der Lehrerbildung. (7. Beiheft der Zeitschrift Die deutsche Schule). Weinheim: Juventa.
- Lehmermöhle, D./Rothgangel, M./Bögelholz, S./Hasselhorn, M./Watermann, R. (Hrsg.) (2007): professionell lehren - erfolgreich lernen. Münster: Waxmann.
- Lüders, M./Wissinger, J. (Hrsg.) (2007): Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation. Münster: Waxmann.
- Mitchell, K.J./Robinson, D.Z./Plake, B.S./Knowles, K.T. (Hrsg.) (2001): Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality. Washington: National Academy Press.
- Peterson, K.D. (2000): Teacher Evaluation. A Comprehensive Guide to New Directions and Practices. (2nd edition). Thousand Oaks: Corwin.
- Porter, A.C./Youngs, P./Odden, A. (2001): Advances in Teacher Assessment and their Uses. In: Richardson, V. (Hrsg.): Handbook of Research on Teaching (4th edition). Washington D.C.: American Educational Research Association, S. 259–297.
- Reintjes, C. (2007): Erziehungswissenschaft – ein notwendiger Bestandteil der gymnasialen Lehrerausbildung? Eine explorative Studie mit Hauptseminarleitern in Nordrhein-Westfalen. Münster: LIT-Verlag.
- Röthland, M./Terhart, E. (2009): Forschung zum Lehrerberuf. In: Tippelt, R./Schmidt, B. (Hrsg.): Handbuch Bildungsforschung. 2. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schaefers, C./Terhart, E. (2006): The Participation of Schools in the Recruitment of Teachers: Evaluation new Procedures in Germany. In: European Journal of Teacher Education 29, S. 505–517.
- Schubarth, W./Speck, K./Seidel, A. (Hrsg.) (2007): Endlich Praxis! Die Zweite Phase der Lehrerbildung. Potsdamer Studien zum Referendariat. Frankfurt: Lang Verlag.
- Seifried, J./Abel, J. (Hrsg.) (2006): Empirische Lehrerbildungsforschung. Stand und Perspektiven. Münster: Waxmann.
- Strecke, T. (2007): Die zweite Phase der Lehrerbildung im Spiegel neuerer empirischer Forschungsergebnisse – unter besonderer Berücksichtigung des Lehramts an Berufskollegs. Schriftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das Lehramt an Berufskollegs. Universität Münster.
- Terhart, E. (2007): Erfassung und Beurteilung der beruflichen Kompetenzen von Lehrkräften. In: Lüders, M./Wissinger, J. (Hrsg.): Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation. Münster: Waxmann, S. 37–62.
- Walke, J. (2004): Die Situation des Ausbildungspersonals in der Zweiten Phase der Lehrerbildung. Eine Analyse auf der Basis vorhandener Informationen und Daten. (ZfL-Text Nr. 4). Universität Münster.
- Walke, J. (2007): Die Zweite Phase der Lehrerbildung. Ein Überblick über Stand, Problemlagen und Reformtendenzen. Eine Expertise für den Wissenschaftlichen Beirat des Aktionsprogramms „Neue Wege in der Lehrerbildung“ des Stifterverbandes für die Deutsche Wissenschaft/Mercator-Stiftung. (Schriftenreihe zur Lehrerbildung, Band III). Essen.

- Weimar, D. (2005): Streß und Flow-Erleben. Eine empirische Untersuchung zur Bedeutung von Kognitionen, Emotionen und Motivation bei Lehramtsstudierenden, Referendaren und Lehrern. Berlin: Logos.
- Wernet, A. (2006): „Man kann ja sagen, was man will: es ist ein Lehrer-Schüler-Verhältnis“. Eine fallanalytische Skizze zu Kollegialitätsproblemen im Referendariat. In: Schubarth, W./Pohlenz, P. (Hrsg.): Qualitätsentwicklung und Evaluation in der Lehrerbildung. Die zweite Phase: Das Referendariat. Potsdam: Universitätsverlag, S. 193–207.
- Wernet, A. (2008): Konformismus statt kollegiale Anerkennung: Fallstudien zur Ausbildungskultur im Referendariat am Beispiel von Beurteilungen. Unveröffentlichtes Manuskript.
- Wolf, K./Lichtenstein, G./Stevenson, C. (1997): Portfolios in Teacher Evaluation. In: Stronge, J.H. (Hrsg.): Evaluating Teaching: A Guide to Current Thinking and Best Practice. Thousand Oaks: Sage, S. 193–214.

Abstract: The issue of the possibilities of a reliable assessment of the professional competences of teachers is being discussed controversially not only in the scientific debate but also within teacher associations and among the public. This debate often neglects the fact that an assessment of (prospective) teachers has been taking place for decades: namely within the framework of the second phase of training, the “Referendariat” (teacher training) or teaching practice. Following a short survey of recent research results regarding the present situation in the “second phase”, the authors report on a research project analyzing the assessment patterns and manuals used by instructors (heads of departments, heads of seminars for advanced students) during the “second phase”. On the basis of 201 assessment patterns and 87 additional questionnaires on the development and application of these patterns, a survey is given on the practice of assessment during the “Referendariat” (teacher training). The results show, among other things, that a large part of these instruments is based on a concept of the teaching profession focussing on instruction, meaning that the breadth of the standards for teacher training is not depicted. With every second instrument it remains unclear how the individual act of assessment is actually carried out. In none of the cases was the test-diagnostic quality of the instruments actually verified. All in all, it becomes clear that with regard to the quality of the processes and instruments of assessment applied during the second phase of teacher training there exists an urgent need for further development.

Anschrift der Autoren

Rolf Strietholt, TU Dortmund, Institut für Schulentwicklungsforschung, Vogelpothsweg 78; 44227 Dortmund. E-Mail: strietholt@ifs.uni-dortmund.de.

Prof. Dr. Ewald Terhart, Westfälische Wilhelms-Universität Münster, Institut für Erziehungswissenschaft, Bispinghof 5/6, 48147 Münster.
E-Mail: ewald.terhart@uni-muenster.de.